



Validation clinique et normes préliminaires des sous-tests de lecture et de dictée du *Test d'évaluation du langage écrit québécois*



Clinical Validation and Preliminary Norms for the Reading and Writing Subtests of the *Test d'évaluation du langage écrit québécois* [Québec Evaluation of Written Language Test]

MOTS-CLÉS

TEST D'ÉVALUATION
DU LANGAGE ÉCRIT
QUÉBÉCOIS (TELEQ)

DYSLEXIE

DYSORTHOGRAPHIE

DIAGNOSTIC
DES TROUBLES
D'APPRENTISSAGES EN
LECTURE ET EN ÉCRITURE
SENSIBILITÉ ET SPÉCIFICITÉ

Patricia Laniel
Bruno Gauthier

Patricia Laniel et Bruno
Gauthier

Département de psychologie,
Université de Montréal,
Montréal, QC, CANADA

Abrégé

Les sous-tests « Lecture de mots et de pseudomots » et « Dictée de mots et de pseudomots » du *Test d'évaluation du langage écrit québécois* ont été développés et prévalidés pour permettre l'évaluation de la dyslexie/dysorthographe chez les enfants francophones québécois de la 2^e à la 6^e année primaire (Beaudry et al., 2020; Laniel et al., 2022). La présente étude visait à consolider leur validation auprès d'un échantillon clinique composé d'enfants ayant une dyslexie/dysorthographe et à offrir des normes préliminaires pour ces sous-tests. Les 220 participants âgés de 7 à 12 ans étaient répartis en deux groupes : contrôle ($n = 171$) et clinique ($n = 49$). Une évaluation du langage écrit comprenant les deux sous-tests du *Test d'évaluation du langage écrit québécois* ainsi qu'une brève évaluation cognitive ont été effectuées. Les courbes de caractéristique de performance ont permis de déterminer les qualités diagnostiques (sensibilité et spécificité) des six mesures obtenues aux sous-tests « Lecture de mots et de pseudomots » (précision et vitesse de lecture de mots et de pseudomots) et « Dictée de mots et de pseudomots » (précision de l'écriture de mots et de pseudomots). Des valeurs seuils optimales et des normes préliminaires sont présentées pour chaque mesure, pour utilisation clinique. L'utilisation conjointe des six mesures (en utilisant le critère de dépasser les valeurs seuils optimales pour quatre des mesures ou plus) équivaut à une sensibilité de 87,76 % et une spécificité de 97,66 %. Ces sous-tests s'avèrent donc des outils pertinents pour utilisation dans le cadre d'une démarche diagnostique ou de dépistage de la dyslexie/dysorthographe.

Rédacteur :
Stefano Rezzonico

Rédacteur en chef :
David H. McFarland

Abstract

The subtests *Lecture de mots et de pseudomots* [Reading of Words and Pseudowords] and *Dictée de mots et de pseudomots* [Writing of Words and Pseudowords] of the *Test d'évaluation du langage écrit québécois* [Québec Evaluation of Written Language Test] were created and prevalidated for assessing dyslexia/dysorthographie in Québec French-speaking children in Grades 2 to 6 (Beaudry et al., 2020; Laniel et al., 2022). This study aimed to consolidate the subtests' validity in a clinical sample of children with dyslexia/dysorthographie and to provide preliminary norms for these subtests. Two hundred and twenty participants aged 7 to 12 years were divided into two groups: control ($n = 171$) and clinical ($n = 49$). Written language skills were assessed using the two subtests of the *Test d'évaluation du langage écrit québécois* [Québec Evaluation of Written Language Test]. A brief cognitive assessment was also carried out. Receiver operating characteristic curves were used to determine the diagnostic sensitivity and specificity of the six measures obtained with the subtests *Lecture de mots et de pseudomots* [Reading of Words and Pseudowords] and *Dictée de mots et de pseudomots* [Writing of Words and Pseudowords]: word-reading accuracy, word-reading speed, pseudoword-reading accuracy, pseudoword-reading speed, word-writing accuracy, and word-writing speed. For clinical use, optimal cut-off values and preliminary norms are presented for each measure. When all six measures are used together (with the criterion of exceeding the optimal cut-off values for four or more of the measures), a sensitivity of 87.76% and a specificity of 97.66% are found. These subtests are therefore relevant tools to use to diagnose or screen children for dyslexia/dysorthographie.

La dyslexie est un trouble neurodéveloppemental caractérisé par des difficultés dans la reconnaissance exacte et fluide des mots écrits ainsi que par de faibles compétences en décodage et en orthographe (International Dyslexia Association, 2002). Ce trouble toucherait de 3 à 5 % des enfants (Institut national de la santé et de la recherche médicale, 2007), mais la prévalence varie grandement selon les critères diagnostiques utilisés et pourrait aller jusqu'à 17,5 % selon les études (Shaywitz et al., 1994). La majorité, voire la totalité, des enfants ayant une dyslexie présentent également des difficultés au niveau de l'acquisition et de la maîtrise de l'orthographe (dysorthographe; Lefebvre et Stanké, 2016), ces difficultés en lecture et en écriture étant regroupés sous l'appellation « Trouble spécifique des apprentissages » dans le Manuel diagnostique et statistique des troubles mentaux (American Psychiatric Association, 2015). La dyslexie entraîne de nombreuses conséquences secondaires, incluant des difficultés de compréhension de lecture et une diminution du temps de lecture nuisant au développement du vocabulaire et des connaissances générales (Lyon et al., 2003). Les élèves ayant une dyslexie sont aussi plus à risque d'échec et d'abandon scolaires et sont également plus nombreux que leurs pairs normolecteurs à souffrir d'anxiété, de symptômes dépressifs et de problèmes d'estime de soi (Daniel et al., 2006; Humphrey et Mullins, 2002; Prior et al., 1999; Rousseau, 2016). Le diagnostic précoce de la dyslexie/dysorthographe, dans lequel les orthophonistes jouent un rôle primordial (American Speech-Language-Hearing Association, 2001), permet toutefois d'en réduire ses impacts négatifs par l'accès à des mesures d'accommodation et de soutien qui favorisent la réussite scolaire, le bien-être et la confiance en soi de l'enfant (Livingston et al., 2018).

Le diagnostic de la dyslexie se fait en utilisant diverses sources d'information, incluant la performance à des épreuves psychométriques standardisées de lecture orale et de dictée de mots (Stanké et Lefebvre, 2016; St-Pierre et al., 2010). Ces outils doivent mesurer à la fois la précision et la vitesse de lecture ainsi que le fonctionnement des deux voies de la lecture experte (phonologique et lexicale; Coltheart et al., 2001; Lefebvre et Stanké, 2016; Pourcin et al., 2016; Siegel et Hurford, 2019). La voie phonologique implique le décodage des mots en faisant correspondre chaque graphème à un phonème puis en les assemblant, et s'évalue par la lecture de pseudomots (mots n'appartenant pas au lexique de la langue, p. ex. « paton »; Casalis, 2005). La voie lexicale implique quant à elle la reconnaissance des mots contenus dans le lexique orthographique à partir du traitement simultané de la séquence des lettres le constituant, et s'évalue par la lecture de mots irréguliers

(qui contiennent des correspondances graphèmes-phonèmes exceptionnelles, comme dans « monsieur », ou rares, comme dans « technicien »; Casalis, 2005; Lefebvre et Stanké, 2016; St-Pierre et al., 2010). Une difficulté marquée à lire/écrire les pseudomots serait révélatrice d'un trouble au niveau de la voie phonologique alors qu'une difficulté à lire/écrire les mots irréguliers relèverait plutôt d'un déficit de la voie lexicale (Lefebvre et Stanké, 2016; Mousty et Alegria, 1999). Différents types de dyslexie existent selon l'atteinte identifiée (dyslexie phonologique ou lexicale; Castles et Coltheart, 1993), bien que l'atteinte des deux voies soit la plus fréquente (dyslexie mixte; Sprenger-Charolles et al., 2001). Dans la conclusion orthophonique terminant l'évaluation, il est d'ailleurs suggéré de préciser s'il y a présence d'une difficulté plus saillante avec l'une ou l'autre des voies (St-Pierre et al., 2010).

Jusqu'à récemment, il n'existait pas au Québec d'épreuves adéquates offertes en français pour évaluer les habiletés de lecture et d'orthographe de mots/pseudomots, certaines n'étant pas adaptées à la clientèle franco-québécoise (sur le plan du choix des mots ou de l'échantillon normatif), d'autres n'évaluant pas les composantes essentielles de la lecture (vitesse et précision en lecture; intégrité des voies phonologique et lexicale), et d'autres ne présentant pas de propriétés psychométriques satisfaisantes (pour une revue détaillée, voir Beaudry et al., 2020; Laniel et al., 2022). Devant l'absence d'outils adéquats, le *Test d'évaluation du langage écrit québécois* (TELEQ) a été créé. Cet outil en développement continu permet jusqu'à présent d'évaluer l'intégrité des voies lexicale et phonologique avec les sous-tests « Lecture de mots et de pseudomots » et « Dictée de mots et de pseudomots ». Ces sous-tests ont fait l'objet d'une prévalidation auprès d'enfants normolecteurs où la fidélité test-retest, la consistance interne et la validité concordante ont été décrites (Beaudry et al., 2020; Laniel et al., 2022).

Les sous-tests « Lecture de mots et de pseudomots » et « Dictée de mots et de pseudomots » ont été développés dans le but d'être utilisés pour le dépistage (identifier les facteurs de risques ou les signes prédictifs du trouble) et le diagnostic (confirmer la présence du trouble et spécifier sa nature; Institut national de la santé et de la recherche médicale, 2007) de la dyslexie par les professionnels habilités à le faire. Pour qu'un outil soit utilisé dans le cadre d'une démarche diagnostique, la présence d'une validité et d'une fidélité adéquates est un critère essentiel, mais non suffisant. Le critère le plus important à considérer par les cliniciens dans la sélection d'un outil diagnostique serait le pouvoir discriminant de celui-ci pour distinguer les enfants avec et sans trouble (Friberg, 2010; Leclercq et Veys,

2014; Plante et Vance, 1994; Spaulding et al., 2006), seul critère permettant de connaître la précision du diagnostic effectué avec cet outil. Afin de démontrer la pertinence de ces sous-tests pour le diagnostic de la dyslexie, il était donc essentiel de démontrer leur sensibilité (capacité à identifier correctement les enfants ayant une dyslexie ou pourcentage de vrais positifs) et leur spécificité (capacité à identifier correctement les enfants n'ayant pas de dyslexie ou pourcentage de vrais négatifs; Lalkhen et McCluskey, 2008) à la dyslexie chez les enfants du primaire fréquentant une école francophone du Québec. Par ailleurs, un test de dépistage doit minimiser les faux négatifs (optimiser la sensibilité) afin de ne pas négliger un enfant à risque alors qu'un test diagnostique doit éviter les faux positifs et donc, optimiser la spécificité (Bertrand et al., 2010). Comme ces deux sous-tests du TELEQ pourront être utilisés à la fois pour le dépistage et le diagnostic de la dyslexie, une optimisation de ces deux critères a été privilégiée. Enfin, pour le clinicien, les valeurs prédictives positive et négative sont aussi, sinon plus pertinentes à la décision clinique que les valeurs de sensibilité et de spécificité d'un test. Les valeurs prédictives permettent d'évaluer la probabilité que l'individu présente ou non la condition en fonction du résultat du test (Delacour et al., 2009). Ainsi, la valeur prédictive positive (VPP) correspond à la probabilité que l'enfant ait effectivement une dyslexie quand le résultat obtenu au test est positif et la valeur prédictive négative (VPN) correspond à la probabilité que l'enfant soit normolecteur quand le résultat au test est négatif. Ces deux valeurs sont fonction de la prévalence de la condition dans la population où le test est utilisé (Delacour et al., 2009; Glaros et Kline, 1988).

Selon McCauley et Swisher (1984), qui ont défini des critères de qualité des tests standardisés ayant été repris par plusieurs ensuite (Bouchard et al., 2009; Friberg, 2010; Leclercq et Veys, 2014; Plante et Vance, 1994), la présence d'un échantillon normatif satisfaisant est un autre critère essentiel à l'utilisation d'un test normalisé. En effet, pour qu'un outil puisse être utilisé dans le cadre d'une démarche diagnostique, il est essentiel de disposer de normes culturelles et linguistiques appropriées pour l'enfant (Bouchard et al., 2009; Lefebvre et Trudeau, 2005).

Ainsi, l'objectif général de la présente étude était de consolider la validation des sous-tests « Lecture de mots et de pseudomots » et « Dictée de mots et de pseudomots » du TELEQ auprès d'un échantillon clinique composé d'enfants ayant reçu un diagnostic de dyslexie/dysorthographe. Pour ce faire, les courbes ROC (de l'anglais *Receiver Operating Characteristic*) des six mesures obtenues avec ces sous-tests et de l'utilisation conjointe

de ces mesures ont été établies afin de déterminer leur sensibilité et leur spécificité ainsi que leurs valeurs prédictives positive et négative pour le diagnostic de la dyslexie/dysorthographe chez les enfants franco-québécois de la 2^e à la 6^e année. Le second objectif poursuivi était de fournir des normes préliminaires pour ces deux sous-tests.

Méthodologie

Participants

Un total de 249 enfants de la 2^e à la 6^e année du primaire ont participé à l'étude. Pour être admissibles, les enfants devaient être scolarisés en français depuis le début de leur scolarité et parler le français à la maison.

Parmi les participants, 200 enfants normolecteurs ont été recrutés pour former le groupe contrôle. De ceux-ci, 128 provenaient d'une étude de prévalidation des sous-tests « Lecture de mots et pseudomots » et « Dictée de mots et de pseudomots » du TELEQ (Étude 1; Beaudry et al., 2020; Laniel et al., 2022) alors que 72 provenaient d'une étude de validation du sous-test « Compréhension de lecture » du TELEQ (Étude 2; Picotte-Lavoie, 2021). Les critères d'exclusion étaient les suivants : avoir un échec scolaire en français (note en deçà de 60 % au bulletin), avoir redoublé une année scolaire ou bénéficié de suivi en orthophonie ou en orthopédagogie en lien avec le langage écrit, avoir subi un traumatisme craniocérébral ou avoir reçu un diagnostic de trouble neurodéveloppemental (trouble de la communication, incluant le trouble développemental du langage et le trouble développemental des sons de la parole; trouble du spectre de l'autisme; trouble du déficit de l'attention/hyperactivité [TDAH]; trouble d'apprentissage; retard intellectuel ou trouble de tics – voir American Psychiatric Association, 2015). De plus, les participants qui ont obtenu un score situé à plus de deux écarts-types sous la moyenne pour leur âge à l'indice de précision ou de vitesse de lecture de l'*Alouette-R* (Lefavrais, 2005), un test couramment utilisé pour le diagnostic de la dyslexie (Bertrand et al., 2010; Cavalli et Colé, 2018), ont été exclus du groupe contrôle. De la même façon, les participants qui ont obtenu un score situé à plus de deux écarts-types sous la moyenne aux mesures d'estimation du quotient intellectuel (mesuré avec la 4^e édition ou la 5^e édition de l'*Échelle d'intelligence de Wechsler pour enfants* [WISC-IV ou WISC-V], voir la section « Mesures ») et ceux ayant obtenu un score situé à plus de deux écarts-types de la moyenne aux symptômes d'inattention ou d'hyperactivité/impulsivité mesurés avec le *ADHD Rating Scale IV* (Étude 1; DuPaul et al., 1998) ou le *Conners – Troisième édition* (Étude 2; Conners, 2008) ont été exclus du groupe contrôle. Ces critères ont été appliqués afin d'exclure les enfants chez qui un trouble neurodéveloppemental non diagnostiqué ayant

un impact sur l'acquisition du langage écrit pouvait être fortement suspecté.

Quarante-neuf enfants ont été recrutés pour former le groupe clinique et ont participé à l'Étude 1. Pour être inclus dans le groupe clinique, les enfants devaient avoir reçu un diagnostic officiel de dyslexie/dysorthographe posé par un professionnel habilité à le faire (psychologue/neuropsychologue ou orthophoniste). Les enfants du groupe clinique ayant un trouble neurologique, une déficience intellectuelle, un trouble du spectre de l'autisme ou ayant déjà subi un traumatisme craniocérébral n'étaient pas admissibles à l'étude.

Procédure

Le présent projet a été approuvé par le comité d'éthique de la recherche en arts et sciences de l'Université de Montréal (n° 2015-16). Les participants ont été recrutés dans des écoles primaires et des camps de jour sur une base volontaire. Les parents qui souhaitaient faire participer leur enfant à l'étude ont fourni leurs coordonnées puis ont été contactés par les chercheurs pour la vérification des critères d'inclusion et d'exclusion et la prise de rendez-vous. La durée de la rencontre était d'environ deux heures pour l'Étude 1 et de trois heures pour l'Étude 2. L'évaluation comprenait une estimation du quotient intellectuel verbal et non verbal, ainsi qu'un examen du langage écrit comprenant le TELEQ et l'*Alouette-R*. Dans les deux études, le premier test réalisé avec l'enfant était le TELEQ. Les parents remplissaient un questionnaire sociodémographique et un questionnaire de dépistage des symptômes du TDAH chez leur enfant. Les rencontres ont eu lieu dans les locaux de l'Université de Montréal, de la commission scolaire Pierre-Neveu, d'une clinique privée en Haute-Yamaska ou encore au domicile des parents. Les évaluations ont eu lieu entre mars 2016 et mars 2020. Elles ont été réalisées par des étudiantes au doctorat en neuropsychologie ou par des étudiantes du baccalauréat en psychologie formées pour l'administration des tests. L'enfant et l'évaluatrice disposaient d'un endroit calme pour procéder à l'évaluation. Tous les dossiers des participants ont été doublement cotés par une des chercheuses afin d'assurer une cotation exacte.

Mesures

TELEQ

Le TELEQ est un nouveau test développé afin de répondre au manque d'outils adaptés aux enfants québécois pour l'évaluation de la dyslexie (Bouchard et al., 2009). Le TELEQ comporte pour l'instant trois sous-tests : « Lecture de mots et de pseudomots », « Dictée de mots

et de pseudomots » et « Compréhension de lecture ». Des sous-tests évaluant les habiletés de conscience morphologique et phonologique sont en cours de développement. Dans le cadre de la présente étude, tous les enfants ont passé deux sous-tests du TELEQ, soit le sous-test « Lecture de mots et de pseudomots » et le sous-test « Dictée de mots et de pseudomots ». Ces sous-tests, destinés aux enfants de la 2^e à la 6^e année primaire, sont basés sur le modèle à double voie de Coltheart (Coltheart et al., 2001) et permettent de détailler le fonctionnement des deux voies de la lecture et de l'écriture. Les données préliminaires de validation de ces sous-tests, ayant été obtenues auprès de 119 enfants franco-québécois normolecteurs de la 2^e à la 6^e année primaire, sont très satisfaisantes. Les corrélations entre les passations reflètent une très bonne fidélité test-retest, la consistance interne des listes est adéquate et la validité concordante est satisfaisante (corrélations fortes avec les mesures de la *Batterie analytique du langage écrit* et de l'*Alouette-R*; Beaudry et al., 2020; Jacquier-Roux et al., 2010; Laniel et al., 2022; Lefavrais, 2005).

Sous-test « Lecture de mots et de pseudomots ».

Ce sous-test mesure la précision et la vitesse de lecture de mots et de pseudomots. Il est composé de 39 mots irréguliers ou inconsistants formant la liste de mots irréguliers et de 40 pseudomots formant la liste de pseudomots. Les items de la liste de mots irréguliers sont tirés de la liste orthographique du Programme de formation de l'école québécoise fournie par le ministère de l'Éducation, du Loisir et du Sport (2014). Les pseudomots ont été construits de manière à représenter la plus grande diversité de règles graphophonémiques possibles et se répartissent en deux niveaux de longueur et de complexité (20 pseudomots courts et 20 pseudomots longs; chaque niveau de longueur comprenant 10 items ayant une structure syllabique simple et 10 ayant une structure syllabique complexe). Ce sous-test fournit quatre mesures : la précision et la vitesse de lecture de la liste de mots irréguliers ainsi que la précision et la vitesse de lecture de la liste de pseudomots.

Sous-test « Dictée de mots et de pseudomots ».

Ce sous-test mesure la précision de l'écriture de mots et de pseudomots. Il est composé de 60 mots irréguliers ou inconsistants formant la liste de mots irréguliers et de 24 pseudomots formant la liste de pseudomots. Les items de la liste de mots irréguliers font partie de la liste orthographique fournie par le ministère de l'Éducation, des Loisirs et du Sport (2014) et sont classés par niveau scolaire auquel ils doivent être enseignés (10 par niveau scolaire). Ces mots irréguliers sont administrés selon des règles d'arrêt et de marche arrière faisant en sorte que les items

ne sont pas tous administrés à chaque enfant (voir Beaudry et al., 2020, 2021 pour une description détaillée). Les pseudomots sont ordonnés selon leur niveau de longueur et de complexité et sont tous administrés, sans règle d'arrêt. Tous les mots/pseudomots de la dictée sont enregistrés sur fichiers audios et les enregistrements sont présentés à l'enfant afin de favoriser la standardisation de la passation. Ce sous-test fournit deux mesures; la précision des mots irréguliers et celle des pseudomots.

WISC-IV et WISC-V

Afin d'estimer le quotient intellectuel verbal et non verbal des enfants, deux sous-tests du WISC-IV (Étude 1) ou du WISC-V (Étude 2) ont été utilisés (Wechsler, 2005, 2015). Pour l'estimation du quotient intellectuel verbal, le sous-test « Vocabulaire » de la WISC-IV ou le sous-test « Similitudes » de la WISC-V ont été choisis, puisqu'ils sont les mieux corrélés à l'indice de Compréhension verbale dans leur batterie respective (Wechsler, 2005, 2015). Pour l'estimation du quotient intellectuel non verbal, le sous-test « Matrices » de la WISC-IV (Étude 1) ou de la WISC-V (Étude 2) a été utilisé. Ces mesures ont été utilisées pour mesurer l'équivalence des groupes contrôle et clinique ainsi qu'aux fins d'exclusion. Les participants du groupe contrôle ayant des scores à plus de deux écarts-types sous la moyenne à l'un ou l'autre de ces sous-tests ont été exclus.

Résultats

Description de l'échantillon

Sur les 200 participants du groupe contrôle évalués, 29 ont été exclus de l'étude parce qu'ils avaient obtenu des scores à plus de deux écarts-types sous la moyenne à l'indice de précision ou de vitesse de lecture mesurée avec l'*Alouette-R* ($n = 13$), aux mesures d'estimation du quotient intellectuel verbal ($n = 4$) ou non verbal ($n = 1$) ou aux symptômes d'hyperactivité/impulsivité ou d'inattention (mesurés avec le *ADHD Rating Scale IV* ou le *Conners – Troisième édition*, $n = 11$). Le groupe contrôle final est composé de 171 enfants, 117 provenant de l'Étude 1 et 54 provenant de l'Étude 2, âgés de 7 ans : 3 mois à 12 ans : 8 mois ($M = 9$ ans : 11 mois, $ET = 1$ an : 4 mois).

Le groupe clinique est composé de 49 participants ayant une dyslexie/dysorthographe dont le diagnostic a été posé par un psychologue/neuropsychologue ou un orthophoniste, âgés de 7 ans : 6 mois à 12 ans : 6 mois ($M = 10$ ans : 7 mois, $ET = 1$ an : 3 mois). Les caractéristiques sociodémographiques de l'échantillon sont présentées dans le **tableau 1**.

Les enfants du groupe clinique avaient reçu leur diagnostic de zéro à cinq ans avant l'évaluation ($M = 1,6$, $ET = 1,3$). Selon les rapports des professionnels consultés,

36 participants du groupe clinique présentaient une dyslexie/dysorthographe mixte, un présentait une dyslexie/dysorthographe phonologique, trois présentaient une dyslexie/dysorthographe lexicale alors qu'un présentait une dysorthographe lexicale (sans dyslexie mentionnée). Les rapports de quatre des participants ne mentionnaient pas le type de dyslexie ou la voie ou les voies de la lecture atteintes et il n'a pas été possible d'obtenir le rapport de quatre autres participants.

Des enfants présentant un trouble comorbide à la dyslexie/dysorthographe ont été inclus dans le groupe clinique afin de représenter adéquatement la population rencontrée en clinique. En effet, la dyslexie est fréquemment rencontrée en comorbidité avec un autre trouble neurodéveloppemental chez un même enfant (Margari et al., 2013; Willcutt et Pennington, 2000). Ainsi, 24 participants inclus dans le groupe clinique avaient également reçu un diagnostic de TDAH, trois présentaient un trouble de langage, deux présentaient un trouble d'acquisition de la coordination et un avait reçu un diagnostic de dyscalculie. Parmi les 24 participants ayant un TDAH, 16 prenaient une médication pour le TDAH (dont 14 qui l'avaient pris le jour de l'évaluation). Le type de médication utilisée était soit un stimulant à base de méthylphénidate (p. ex. Aderall, Vyvanse, Foquest; $n = 9$), un stimulant à base d'amphétamine (p. ex. Biphentin, Concerta; $n = 5$), un non-stimulant (Intuniv; $n = 1$) ou une combinaison d'un stimulant à base de méthylphénidate et d'un non-stimulant ($n = 1$). Finalement, dix des participants ayant une dyslexie avaient doublé une année scolaire et la majorité des participants du groupe clinique ($n = 43$) recevaient au moment de l'étude des interventions liées à leurs difficultés de lecture/écriture (p. ex. suivi avec un orthopédagogue ou un orthophoniste à l'école).

Analyses préliminaires et statistiques descriptives

Nous avons premièrement examiné l'équivalence des groupes (contrôle et clinique) ainsi que le lien entre certaines variables sociodémographiques et les scores obtenus aux mesures du TELEQ. La normalité des distributions des scores pour l'ensemble des variables a d'abord été vérifiée à l'aide des scores d'asymétrie et de voussure (Curran et al., 1996). Certaines variables n'étaient pas distribuées normalement, ce qui a été considéré dans les analyses subséquentes. Les analyses statistiques ont été effectuées à l'aide des logiciels IBM SPSS Statistics 25 et Medcalc 14. Le seuil critique utilisé était de $p < 0,05$.

L'équivalence des groupes contrôle et clinique a été examinée pour le genre, le plus haut niveau de scolarité complété par la mère, le rang décile de l'indice de milieu

socio-économique de l'école fréquentée ainsi que l'estimation du quotient intellectuel verbal et non verbal des participants. D'abord, un test du khi carré a montré que les groupes clinique et contrôle différaient en ce qui a trait au genre des participants ($\chi^2(1) = 9,27, p = 0,002$). En effet, le groupe contrôle était composé en majorité de filles (57,31 %) alors que le groupe clinique était composé d'une majorité de garçons (67,35 %). Les groupes contrôle et clinique différaient aussi en ce qui a trait au plus haut niveau de scolarité complété par la mère. En effet, un khi carré montre qu'il y avait davantage de mères ayant un diplôme

universitaire (1^{er}, 2^e ou 3^e cycle) dans le groupe contrôle (64,71 %) que dans le groupe clinique (53,06 %; $\chi^2(1) = 5,04, p = 0,025$). Le groupe contrôle et le groupe clinique se répartissaient toutefois équitablement entre les rangs déciles de l'indice de milieux socio-économique de l'école fréquentée « favorisés » (rangs 1 à 5) et « défavorisés » (rangs 6 à 10) comme le suggère un Khi-carré ($\chi^2(1) = 1,31, p = 0,253$). Finalement, des tests *t* indiquent que la moyenne de l'estimation du quotient intellectuel verbal du groupe contrôle ($M = 11,35, ET = 2,76$) était significativement plus grande que celle du groupe

Tableau 1		
Caractéristiques sociodémographiques de l'échantillon		
Variables	Groupe contrôle	Groupe clinique
	(n = 171)	(n = 49)
	n	n
Sexe		
Filles	98	16
Garçons	73	33
Niveau scolaire		
2 ^e année	25	2
3 ^e année	32	10
4 ^e année	41	12
5 ^e année	45	10
6 ^e année	28	15
Région de l'école fréquentée		
Laval	97	21
Montérégie	31	8
Montréal	26	15
Laurentides	14	3
Ville de Québec	2	1
Lanaudière	1	1
Revenu familial annuel brut		
39 000 \$ et moins	9	7
40 000 à 59 000 \$	19	4
60 000 à 79 000 \$	17	8
80 000 à 99 000 \$	24	10
100 000 \$ et plus	101	19
Information non disponible	1	1
Rang décile de l'Indice de milieu socio-économique de l'école fréquentée^a		
1 à 5	98	21
6 à 10	66	20
Information non disponible	7	8

Tableau 1 (suite)		
Caractéristiques sociodémographiques de l'échantillon		
Variables	Groupe contrôle	Groupe clinique
	(<i>n</i> = 171)	(<i>n</i> = 49)
	<i>n</i>	<i>n</i>
Plus haut niveau de scolarité atteint par la mère		
Aucun	1	1
D.E.S.	9	7
D.E.P.	14	9
Diplôme collégial (D.E.C. ou A.E.C.)	36	9
Diplôme de 1 ^{er} cycle universitaire (Baccalauréat ou certificat)	81	10
Diplôme de 2 ^e ou 3 ^e cycle universitaire (D.E.S.S., Maîtrise ou Doctorat)	29	13
Information non disponible	1	0
Langue parlée à la maison en plus du français		
Aucune	133	43
Arabe	11	2
Anglais	10	2
Espagnol	5	1
Créole	5	0
Berbère ou Kabyle	3	0
Russe	2	0
Libanais	1	0
Chinois	1	0
Anglais et espagnol	0	1

Note. D.E.S. = Diplôme d'études secondaires; D.E.P. = Diplôme d'études professionnelles; D.E.C. = Diplôme d'études collégiales; A.E.C. = Attestation d'études collégiales; D.E.S.S. = Diplôme d'études supérieures spécialisées.

*Le rang décile de l'Indice de milieu socio-économique de l'école fréquentée est un indice de défavorisation fourni par le ministère de l'Éducation et de l'Enseignement supérieur. Les écoles sont classées sur une échelle allant de 1 à 10, le rang 1 étant considéré comme le moins défavorisé et le rang 10 comme le plus défavorisé (ministère de l'Éducation et de l'Enseignement supérieur, 2020).

clinique ($M = 9,22$, $ET = 3,01$, $t(215) = 4,56$, $p < 0,001$), tout comme la moyenne de l'estimation du quotient intellectuel non verbal du groupe contrôle ($M = 10,99$, $ET = 2,77$) relativement au groupe clinique ($M = 9,62$, $ET = 2,33$, $t(216) = 3,12$, $p < 0,001$).

Ensuite, l'association entre les variables sociodémographiques et les scores au TELEQ a été mesurée, en utilisant les scores Z au TELEQ (moyenne et écart-type du groupe contrôle par niveau scolaire : $\text{Score} - M/ET$). Des analyses de variance ayant pour facteur le genre ont montré que, dans les deux groupes, les scores des filles aux six mesures du TELEQ ne différaient pas significativement de ceux des garçons (voir le **tableau 2**). Toutefois, des analyses de variance ayant pour facteur le plus haut niveau de scolarité complété par la mère (universitaire ou non-universitaire) ont montré que, dans le groupe contrôle, les scores des enfants dont la mère détient un diplôme universitaire obtenaient

en moyenne un score plus élevé que ceux dont la mère ne détient pas de diplôme universitaire aux mesures de précision de lecture des mots irréguliers, de vitesse de lecture des pseudomots et aux deux dictées du TELEQ (voir le **tableau 3**). La taille de ces effets était toutefois petite à moyenne selon les critères de Cohen (1988).

Des corrélations positives moyennes étaient présentes entre l'estimation du quotient intellectuel verbal et le score Z des enfants à la précision de lecture de mots irréguliers, autant dans le groupe contrôle ($r = 0,23$, $p = 0,003$) que dans le groupe clinique ($r = 0,34$, $p = 0,023$; Cohen, 1988). Dans le groupe contrôle, l'estimation du quotient intellectuel verbal était également reliée au score Z de la vitesse de lecture des pseudomots ($r = -0,17$, $p = 0,024$), de sorte que plus les enfants avaient un quotient intellectuel verbal élevé, plus ils lisaient les listes rapidement, et au score Z de la dictée de pseudomots ($r = 0,18$, $p = 0,021$).

Tableau 2								
Effet du genre sur les mesures du TELEQ								
	Groupe contrôle				Groupe clinique			
	Filles (n = 98)	Garçons (n = 73)	F	p	Filles (n = 16)	Garçons (n = 33)	F	p
Lecture de mots irréguliers – Précision	0,07 (0,95)	-0,09 (1,04)	1,14 ^a	0,29	-3,08 (2,65)	-2,11 (1,96)	2,04 ^b	0,16
Lecture de mots irréguliers – Vitesse ^c	-0,11 (1,04)	0,16 (0,89)	3,22 ^a	0,07	-4,67 (4,59)	-4,46 (5,85)	0,02 ^b	0,90
Lecture de pseudomots – Précision	0,04 (1,03)	-0,06 (0,94)	0,46 ^a	0,50	-2,62 (1,96)	-2,13 (1,59)	0,87 ^b	0,36
Lecture de pseudomots – Vitesse ^c	-0,08 (1,05)	0,11 (0,90)	1,50 ^a	0,23	-3,61 (4,22)	-2,98 (3,85)	0,27 ^b	0,60
Dictée de mots irréguliers	0,03 (1,00)	-0,049 (0,97)	0,19 ^d	0,67	-3,48 (2,38)	-3,54 (2,69)	0,01 ^e	0,94
Dictée de pseudomots	-0,01 (1,02)	0,02 (0,95)	0,05 ^a	0,83	-1,60 (1,33)	-1,38 (1,21)	0,33 ^b	0,57

Note. Les moyennes (et écart-types) sont présentées dans les colonnes « Filles » et « Garçons ». Des analyses non paramétriques (*U* de Mann-Whitney) ont également été effectuées. Comme les résultats sont les mêmes que pour les analyses de variance, seules celles-ci sont présentées. TELEQ = Test d'évaluation du langage écrit québécois; *ddl* = degrés de liberté.

^a*ddl* = 1, 169. ^b*ddl* = 1, 47. ^cLes scores *Z* des indices de vitesse ont été inversés afin que les enfants dont les temps de lecture sont les plus longs obtiennent des scores *Z* sous la moyenne de leur âge. ^d*ddl* = 1, 155. ^e*ddl* = 1, 45.

Tableau 3									
Effet du plus haut niveau de scolarité complété par la mère sur les mesures du TELEQ									
	Groupe contrôle					Groupe clinique			
	Mère n'ayant pas de diplôme universitaire (n = 60)	Mère ayant un diplôme universitaire (n = 110)	F	p	R ²	Mère n'ayant pas de diplôme universitaire (n = 26)	Mère ayant un diplôme universitaire (n = 23)	F	p
Lecture de mots irréguliers – Précision	-0,26 (1,02)	0,15 (0,95)	6,90 ^a	0,01	0,04	-2,90 (2,59)	-1,90 (1,63)	2,55 ^b	0,12
Lecture de mots irréguliers – Vitesse ^c	-0,16 (1,22)	0,09 (0,83)	2,44 ^a	0,17	-	-5,62 (6,54)	-3,30 (3,57)	2,28 ^b	0,14
Lecture de pseudomots – Précision	-0,17 (1,02)	0,09 (0,97)	2,58 ^a	0,11	-	-2,65 (1,90)	-1,89 (1,40)	2,50 ^b	0,12
Lecture de pseudomots – Vitesse ^c	-0,22 (0,99)	0,13 (0,97)	4,97 ^a	0,03	0,03	-3,94 (4,96)	-2,33 (2,09)	2,10 ^b	0,15
Dictée de mots irréguliers	-0,27 (1,01)	0,14 (0,95)	6,35 ^d	0,01	0,04	-3,16 (2,26)	-3,94 (2,88)	0,07 ^e	0,31
Dictée de pseudomots	-0,30 (0,96)	0,16 (0,98)	8,66 ^a	0,004	0,05	-1,74 (1,14)	-1,13 (1,30)	3,05 ^b	0,09

Notes. Les moyennes (et écart-types) sont présentées dans les colonnes « Mère n'ayant pas de diplôme universitaire » et « Mère ayant un diplôme universitaire ». Des analyses non paramétriques (*U* de Mann-Whitney) ont également été effectuées. Comme les résultats sont les mêmes que pour les analyses de variance, seules celles-ci sont présentées. TELEQ = Test d'évaluation du langage écrit québécois; *ddl* = degrés de liberté.

^a*ddl* = 1, 168. ^b*ddl* = 1, 47. ^cLes scores *Z* des indices de vitesse ont été inversés afin que les enfants dont les temps de lecture sont les plus longs obtiennent des scores *Z* sous la moyenne de leur âge. ^d*ddl* = 1, 154. ^e*ddl* = 1, 45.

Toujours dans le groupe contrôle, l'estimation de quotient intellectuel non verbal était positivement reliée au score à la dictée de pseudomots ($r = 0,21, p = 0,006$). Les autres corrélations entre les estimations du quotient intellectuel et les mesures du TELEQ étaient non-significatives ($p > 0,05$).

Comparaison des scores entre les groupes

Des analyses de variance ont été réalisées pour comparer les groupes contrôle et clinique par niveau scolaire (3^e à 6^e année) sur chacune des mesures du TELEQ et sont présentées dans le **tableau 4**. Les enfants de 2^e année

n'étaient pas inclus dans ces comparaisons en raison du trop petit échantillon de participants de 2^e année ayant une dyslexie/dysorthographe ($n = 2$, dont un qui n'a pas pu terminer la lecture de la liste de pseudomots). La correction de Benjamini-Hochberg a été appliquée aux 24 valeurs de p obtenues afin de contrôler le *False Discovery rate* (Benjamini et Hochberg, 2000). Toutes les différences étaient statistiquement significatives et avaient une grande taille d'effet selon le critère de Cohen (1988) de $R^2 \geq 0,14$ (voir le **tableau 4**), sauf à la dictée de pseudomots en 6^e année où l'effet est considéré de taille moyenne.

Tableau 4

Comparaisons des scores au TELEQ entre le groupe contrôle et le groupe clinique, par niveau scolaire

	3 ^e année				4 ^e année				5 ^e année				6 ^e année			
	GCO (n = 32)	GCL (n = 10)	F ^a	R ²	GCO (n = 41)	GCL (n = 12)	F	R ²	GCO (n = 45)	GCL (n = 10)	F	R ²	GCO (n = 28)	GCL (n = 15)	F ^f	R ²
LIR Précision (sur 39)	35,72 (2,77)	24,30 (5,29)	81,00***	0,67	35,32 (2,86)	28,58 (4,89)	36,40 ^b ***	0,42	36,96 (1,89)	34,80 (2,25)	9,90 ^d ***	0,16	37,89 (1,45)	35,13 (3,85)	11,53**	0,22
LIR Vitesse (sec)	37,84 (10,64)	116,74 (63,89)	47,13***	0,54	34,32 (9,72)	97,70 (70,87)	32,23 ^b ***	0,39	28,32 (6,69)	42,09 (8,86)	30,72 ^d ***	0,37	26,21 (6,99)	39,41 (15,34)	15,13***	0,27
LPS Précision (sur 40)	31,47 (3,58)	21,20 (6,25)	42,87***	0,52	30,88 (4,08)	21,33 (6,35)	38,82 ^b ***	0,43	31,62 (3,72)	26,80 (6,49)	10,20 ^d ***	0,16	33,07 (3,10)	26,00 (5,17)	31,59***	0,44
LPS Vitesse (sec)	81,55 (16,48)	159,50 (79,21)	28,54***	0,42	79,22 (14,66)	143,01 (84,87)	21,93 ^b ***	0,30	68,84 (14,5)	93,02 (20,65)	19,39 ^d ***	0,27	60,74 (15,27)	87,47 (19,46)	24,67***	0,38
DIR (sur 60)	31,18 (11,14)	7,70 (5,66)	40,21***	0,53	35,53 (11,37)	12,82 (10,40)	35,26 ^c ***	0,43	46,44 (8,16)	25,22 (9,77)	47,06 ^e ***	0,49	55,11 (4,1)	28,93 (10,94)	128,82***	0,76
DPS (sur 24)	17,72 (3,31)	11,80 (3,68)	23,11***	0,37	18,24 (3,67)	11,50 (3,73)	31,09 ^b ***	0,38	19,47 (2,49)	15,70 (4,24)	14,13 ^d ***	0,21	19,14 (2,8)	16,87 (2,92)	6,27*	0,13

Note. Les moyennes (et écart-types) sont présentées dans les colonnes GCO et GCL. Les valeurs de p ont été corrigées selon la méthode de Benjamini-Hochberg. TELEQ = Test d'évaluation du langage écrit québécois; LIR = Lecture de mots irréguliers; LPS = Lecture de pseudomots; DIR = Dictée de mots irréguliers; DPS = Dictée de pseudomots; sec = secondes. GCO = Groupe contrôle; GCL = Groupe clinique; ddl = degrés de liberté.

^addl = 1,40. ^bddl = 1,51. ^cddl = 1,47. ^dddl = 1,53. ^eddl = 1,50. ^fddl = 1,41.

*** $p < 0,001$. ** $p < 0,01$. * $p < 0,05$.

Sensibilité et spécificité

Des courbes ROC ont été effectuées pour chaque mesure du TELEQ, par niveau scolaire. La courbe ROC représente la relation entre la sensibilité et la spécificité d'un test, pour chaque valeur seuil possible. Le choix de la valeur seuil affecte les valeurs de sensibilité et de spécificité du diagnostic (Morin et al., 1998). Le test idéal aurait une sensibilité et une spécificité de 100 % et classerait ainsi tous les participants dans le bon groupe, sans erreur. Toutefois, en réalité, un test n'est jamais parfait. La meilleure valeur seuil est donc celle qui permet de classer le plus de participants dans le bon groupe, en minimisant les faux positifs et les faux négatifs. Le choix de cette valeur seuil résulte d'un compromis entre sensibilité et spécificité (Bertrand et al., 2010; Morin et al., 1998).

Les scores des enfants du groupe clinique qui avaient une dyslexie/dysorthographe mixte ont été conservés pour effectuer les courbes ROC des six mesures du TELEQ. Les scores des enfants du groupe clinique qui avaient une dyslexie/dysorthographe lexicale ($n = 4$) ont été conservés seulement pour le calcul des courbes ROC des mesures associées aux listes de mots irréguliers. Les scores de l'enfant qui avait une dyslexie/dysorthographe phonologique ($n = 1$) ont été conservés seulement pour le calcul des courbes ROC des mesures associées aux listes de pseudomots. Pour les enfants dont le type de dyslexie n'était pas disponible ($n = 8$), une analyse de la voie ou des voies atteintes a été réalisée en calculant les scores Z de l'enfant pour chaque mesure du TELEQ (moyenne et écart-type du groupe contrôle par niveau scolaire : Score - M/ET). Si le score Z se situait à plus de 1,5 ET sous la moyenne de son niveau scolaire pour au moins une mesure évaluant la voie (lexicale ou phonologique), cette voie était jugée comme atteinte. Ceci nous a permis de classer les enfants pour lesquels le type de dyslexie n'était pas précisé dans le groupe dyslexie mixte ($n = 7$) ou dyslexie phonologique ($n = 1$). Les valeurs seuils associées aux indices de Youden (1950) ont été sélectionnées comme seuil puisqu'elles maximisent la sensibilité et la spécificité (Perkins et Schisterman, 2006). Celles-ci sont présentées dans le **tableau 5**, avec leurs sensibilité et spécificité associées.

Pour mesurer la sensibilité et la spécificité du TELEQ en entier, en incluant les six mesures obtenues aux deux sous-tests, nous avons calculé combien de seuils chaque enfant dépassait sur les six mesures. Une variable a été créée, soit le nombre de seuils dépassés par enfant pour son niveau scolaire, variant de 0 à 6. Nous avons ensuite calculé la sensibilité et la spécificité de cette variable à l'aide d'une courbe ROC. Cette technique attribue autant de poids à chacune des mesures et a été privilégiée puisqu'elle

est directement applicable par les cliniciens dans leur pratique et en raison du grand nombre de mesures (six) qui empêchait l'utilisation de la régression logistique (Bujang et al., 2018). L'aire sous la courbe ROC pour le nombre de seuils dépassés sur les six mesures du TELEQ a ensuite été obtenue. L'aire sous la courbe ROC fournit une mesure globale de la performance d'un test pour distinguer entre deux groupes. Elle varie entre 0,5 (test dont la probabilité de classer les gens dans le bon groupe relèverait du hasard) et 1,0 (test parfait qui classerait les gens dans le bon groupe dans 100 % des cas; Hajian-Tilaki, 2013). L'aire sous la courbe ROC pour le nombre de seuils dépassés sur les six mesures du TELEQ était de 0,98, (avec un intervalle de confiance à 95 % [0,96; 1,00]). Lorsque la valeur seuil optimale déterminée par l'indice de Youden est utilisée comme critère (≥ 4 seuils dépassés), ceci équivaut à une sensibilité de 87,76 % et une spécificité de 97,66 %. D'autres valeurs seuil permettent d'atteindre une sensibilité ou une spécificité maximale, comme présenté dans le **tableau 6**.

Valeurs prédictives positive et négative

Les formules suivantes permettent de calculer les VPP et VPN du TELEQ, en utilisant la sensibilité de 87,76 % et la spécificité de 97,66 %, selon la prévalence dans un contexte donné :

$$VPN = \frac{0,9766 \times (1 - \text{prévalence})}{(0,9766 \times (1 - \text{prévalence})) + ((1 - 0,8776) \times \text{prévalence})}$$

$$VPP = \frac{0,8776 \times \text{prévalence}}{(0,8776 \times \text{prévalence}) + ((1 - 0,9766) \times (1 - \text{prévalence}))}$$

L'utilisation d'une prévalence de 5 % de la dyslexie tirée de la littérature (Institut national de la santé et de la recherche médicale, 2007) équivaut à une VPP de 66,37 % et une VPN de 99,34 % pour le TELEQ, alors que l'utilisation d'une prévalence de 17,5 % de la dyslexie (Shaywitz et al., 1994) équivaut à une VPP de 88,8 % et une VPN de 97,4 %.

Normes préliminaires

Les normes préliminaires ont été obtenues auprès des 171 enfants du groupe contrôle. Comme aucune différence significative entre les garçons et les filles n'a été observée aux scores du TELEQ (voir plus haut la section « Analyses préliminaires »), les normes préliminaires du TELEQ sont présentées par strate de niveau scolaire seulement, dans le **tableau 7**. La présentation des normes sous forme de percentiles a été privilégiée en raison de la distribution de fréquences des scores de vitesse de lecture de mots irréguliers qui ne suit pas une courbe normale (empêchant l'utilisation de scores Z; Aguert et Capel, 2018).

Tableau 5						
Valeurs seuils optimales et sensibilité/spécificité associées, par niveau scolaire, pour chaque mesure du TELEQ						
	Lecture de mots irréguliers – Précision (sur 39)	Lecture de mots irréguliers – Vitesse (secondes)	Lecture de pseudomots – Précision (sur 40)	Lecture de pseudomots – Vitesse (secondes)	Dictée de mots irréguliers (sur 60)	Dictée de pseudomots (sur 24)
2^e année						
Valeur seuil ^a	≤ 6	> 176,17	≤ 3	> 165,5	≤ 0	≤ 7
Se	100	100	100	100	100	100
Sp	100	100	100	100	95	92
3^e année						
Valeur seuil ^a	≤ 31	> 53,34	≤ 26	> 92	≤ 12	≤ 14
Se	100	90	90	90	90	80
Sp	87,5	93,75	87,5	81,25	96,43	84,37
4^e année						
Valeur seuil ^a	≤ 33	> 41,7	≤ 25	> 97,04	≤ 15	≤ 14
Se	91,67	91,67	90	70	72,73	100
Sp	80,49	87,8	87,8	92,68	94,74	82,93
5^e année						
Valeur seuil ^a	≤ 37	> 32,97	≤ 27	> 74,8	≤ 34	≤ 17
Se	100	88,89	55,56	88,89	100	66,67
Sp	51,11	86,67	84,44	66,67	93,02	80
6^e année						
Valeur seuil ^a	≤ 37	> 27,72	≤ 31	> 66,94	≤ 36	≤ 18
Se	92,31	92,31	92,86	92,86	92,86	71,43
Sp	75	67,86	75	82,14	100	71,43

Note. TELEQ = Test d'évaluation du langage écrit québécois. Se = Sensibilité. Sp = Spécificité.

^aLes valeurs seuils optimales sont celles associées aux indices de Youden. Ces valeurs seuils ne devraient pas être utilisées en 2^e année compte tenu de la faible taille de l'échantillon du groupe clinique à ce niveau scolaire (n = 2).

Tableau 6			
Sensibilité et spécificité pour différentes valeurs seuils du nombre de seuils dépassé au TELEQ sur les 6 mesures			
	Valeur seuil	Sensibilité	Spécificité
Sensibilité maximale	≥ 2	100	74,27
	≥ 3	95,92	88,30
Seuil optimal	≥ 4	87,76	97,66
	≥ 5	75,51	99,42
Spécificité maximale	= 6	46,94	100,00

Note. Les caractères gras indiquent le seuil optimal déterminé à l'aide de l'indice de Youden. TELEQ = Test d'évaluation du langage écrit québécois.

Tableau 7
Distribution des scores du TELEQ en centiles, par niveau scolaire

	Centiles	Lecture de mots irréguliers – Précision (sur 39)	Lecture de mots irréguliers – Vitesse (secondes) ^a	Lecture de pseudomots – Précision (sur 40)	Lecture de pseudomots – Vitesse (secondes) ^a	Dictée de mots irréguliers (sur 60)	Dictée de pseudomots (sur 24)
2^e année (n = 25)	90	35	37,88	33,5	76,28	19,5	19
	75	33,5	47,26	30,5	91,11	15,5	17
	50	29	67,49	28	106,72	10	14
	25	25,5	86,69	22	130,64	5,5	9
	10	20,5	129,65	18,5	153,00	3,5	7,5
3^e année (n = 32)	90	39	25,52	35	66,72	44,5	22,5
	75	37	30,42	34	68,73	39,5	20,5
	50	36,5	37,87	32	77,77	33,5	17
	25	35	43,39	30	88,26	22	16
	10	30,5	52,12	25,5	109,38	16	14
4^e année (n = 41)	90	38	23,67	35	61,17	47	22
	75	37,5	29,15	34	68,92	45,5	21
	50	36	32,75	32	77,34	39	19
	25	34	38,26	28	91,17	25,5	15,5
	10	31,5	45,37	25	96,86	19	13
5^e année (n = 45)	90	39	20,67	36	49,19	57	22
	75	38	24,08	34	59,32	54	21,5
	50	38	26,85	32	68,81	46	19
	25	36	31,78	29	80,69	39	18
	10	34,5	37,66	26	89,57	35	16
6^e année (n = 28)	90	39	17,86	37	42,79	58,5	22,5
	75	39	20,53	36	52,21	58	21
	50	38	26,02	32,5	58,51	56	20
	25	37,5	30,95	31,5	65,71	54,5	18
	10	35,5	35,93	28	83,74	45,5	13,5

Note. TELEQ = Test d'évaluation du langage écrit québécois.

^aLes rangs centiles des indices de vitesse ont été inversés afin que les enfants dont les temps de lecture sont les plus longs obtiennent les rangs centiles les plus bas.

Discussion

L'objectif de la présente étude était de consolider la validation des sous-tests « Lecture de mots et de pseudomots » et « Dictée de mots et de pseudomots » du TELEQ, destinés à l'évaluation de la dyslexie/dysorthographe chez les enfants francophones du Québec de la 2^e à la 6^e année primaire. Pour cela, deux objectifs spécifiques ont été poursuivis, soit l'établissement de la sensibilité, de la spécificité et des valeurs prédictives positive et négative de ces sous-tests à la dyslexie/dysorthographe et la création

de normes préliminaires. Chacun de ces objectifs sera discuté ci-dessous.

Comparaison des scores entre les groupes

Avant d'évaluer la sensibilité et la spécificité des mesures du TELEQ à la dyslexie chez les enfants de la 2^e à la 6^e année primaire, nous souhaitons d'abord vérifier si le groupe clinique se distinguait significativement du groupe contrôle sur les différentes mesures du TELEQ. Comme attendu, le groupe clinique obtenait en moyenne un score significativement inférieur à celui obtenu par le groupe

contrôle sur toutes les mesures du TELEQ. Pour presque toutes ces comparaisons, la taille d'effet est grande (Cohen, 1988), signifiant que l'appartenance au groupe contrôle ou clinique explique une grande proportion de la variance des performances au TELEQ. La taille d'effet est considérée comme moyenne pour la comparaison des scores des deux groupes à la dictée de pseudomots en 6^e année. Il est possible que cette mesure soit moins sensible que les autres mesures du TELEQ à la dyslexie en 6^e année. Il est aussi probable que plusieurs enfants du groupe clinique de 6^e année aient bénéficié d'interventions orthopédagogiques ciblant la procédure phonologique, diminuant ainsi l'écart entre la performance des deux groupes sur cette mesure.

Sensibilité et spécificité

La courbe ROC du TELEQ a été obtenue en utilisant le nombre de seuils dépassés sur les 6 mesures du TELEQ par enfant, pour son niveau scolaire. Le TELEQ obtient une aire sous la courbe ROC de 0,98, ce qui est très satisfaisant puisqu'une valeur supérieure à 0,9 indique que le TELEQ constitue un outil remarquable pour discriminer entre le groupe contrôle et le groupe clinique (Hosmer et Lemeshow, 2000). Lorsque le critère d'avoir dépassé quatre seuils ou plus (sur six) est utilisé, ceci équivaut à une sensibilité de 87,76 % et une spécificité de 97,66 %. Ainsi, en utilisant ce critère et ces seuils, le TELEQ répondrait aux exigences d'un test de dépistage de la dyslexie. En effet, bien qu'il n'existe pas de critères généralement acceptés pour ces mesures, puisqu'ils devraient dépendre des implications des résultats du test (par exemple, des coûts associés aux faux positifs et aux faux négatifs), certains auteurs soutiennent qu'une sensibilité supérieure à 80 % et une spécificité supérieure à 90 % seraient requises pour ce type de test (Glascoe et Byrne, 1993). Notons que le TELEQ possède une très bonne spécificité. Son utilisation, avec les seuils présentés, conduirait à un résultat négatif chez 97,66 % des normolecteurs, ce qui fait en sorte qu'il est également tout à fait adapté à être utilisé comme outil diagnostique, pour confirmer la présence du trouble (Bertrand et al., 2010). Ainsi, le clinicien habilité à poser le diagnostic de la dyslexie/dysorthographe et souhaitant utiliser le TELEQ pour le faire peut utiliser les valeurs seuils optimales pour le niveau scolaire de l'enfant qu'il évalue (voir le **tableau 5**) afin de vérifier combien de seuils sont rencontrés. Si l'enfant obtient des résultats dépassant les seuils pour au moins 4 mesures du TELEQ, on peut conclure qu'il est probable qu'il présente une dyslexie/dysorthographe. Le clinicien peut également utiliser un seuil différent en se référant au **tableau 6** selon qu'il souhaite optimiser la sensibilité (p. ex. pour un dépistage) ou la spécificité du test (p. ex. pour un diagnostic). À titre d'exemple, un clinicien effectuant un premier dépistage pourrait privilégier un critère de 2 ou de

3 seuils dépassés ou plus, afin d'augmenter la sensibilité. Il pourrait ensuite poursuivre l'évaluation avec les enfants qui atteignent ce critère ou les référer pour une évaluation plus complète. Le clinicien est toutefois invité à consulter aussi les normes préliminaires afin de situer la performance de l'enfant par rapport à celle d'un groupe d'enfants normolecteurs. Également, notons que ces seuils ont été établis à partir des résultats de seulement deux enfants du groupe clinique en 2^e année; ils ne devraient donc pas être utilisés à ce niveau scolaire (on leur privilégiera l'utilisation des normes préliminaires). De plus, soulignons l'importance pour les cliniciens de ne pas se fier à un seul test lors de l'établissement d'un diagnostic de dyslexie. Bien que l'utilisation d'un test standardisé de lecture et de dictée de mots soit une part importante de l'évaluation, le clinicien se doit d'accumuler différentes sources d'information. Les antécédents familiaux et médicaux, la trajectoire développementale du langage oral, le développement socioaffectif, le contexte d'apprentissage et l'analyse qualitative des types d'erreurs en lecture/écriture ne sont que quelques-uns des éléments à prendre en compte lors de l'établissement du diagnostic (St-Pierre et al., 2010).

Valeurs prédictives positive et négative

Les valeurs prédictives positive et négative du TELEQ ont été présentées pour deux prévalences de la dyslexie/dysorthographe situées aux extrêmes de l'étendue retrouvée dans la littérature. Pour une prévalence de 5 %, la VPN était très élevée (99,34 %) comparativement à la VPP (66,37 %). Dans ce contexte, un score négatif au TELEQ (3 mesures ou moins dépassant le seuil) serait une très bonne indication de l'absence de dyslexie/dysorthographe chez un enfant, alors qu'un score positif serait moins indicatif. Toutefois, avec une prévalence plus élevée (p. ex. 17,5 %) les valeurs prédictives obtenues sont différentes (VPN de 97,4 % et VPP de 88,8 %). L'augmentation importante de la VPP lorsque nous la calculons avec une prévalence plus élevée permet d'illustrer l'importance de mesurer la VPP et la VPN selon la prévalence rencontrée dans un milieu clinique donné. Avec les formules fournies dans le présent article, les valeurs prédictives positive et négative pourront être calculées par les cliniciens selon la prévalence de la dyslexie dans leur milieu clinique.

Normes préliminaires

Les normes obtenues auprès de 171 enfants normolecteurs de la 2^e à la 6^e année primaire présentées dans cet article pourront être utilisées autant en clinique qu'en recherche. Elles constituent toutefois des normes préliminaires, puisqu'elles ont été obtenues auprès d'une moyenne de 34 enfants par tranche d'âge, ce qui est en deçà des 100 enfants par tranche d'âge préconisés par

certain auteurs (McCauley et Swisher, 1984). Notons que les normes ont été réalisées auprès d'enfants normolecteurs majoritairement monolingues francophones et qui provenaient en majorité de familles ayant un statut socio-économique élevé. Ainsi, l'échantillon normatif n'est peut-être pas représentatif de toute la population des enfants québécois fréquentant les écoles primaires francophones. En ce sens, le clinicien est appelé à utiliser son jugement clinique lors de l'interprétation des résultats. Par ailleurs, une étude de normalisation à grande échelle devrait être envisagée, tout comme une étude évaluant davantage l'impact des variables sociodémographiques sur les scores au TELEQ.

Limitations

Cette étude comporte certaines limites qu'il importe de mentionner. En premier lieu, les groupes clinique et contrôle n'étaient pas équivalents sur certaines caractéristiques sociodémographiques, soit le genre, le plus haut niveau de scolarité complété par la mère ainsi que l'estimation du quotient intellectuel, ce qui a pu influencer les comparaisons de groupes et les mesures de sensibilité et de spécificité du test. D'abord, le groupe clinique était constitué d'une majorité de garçons alors que le groupe contrôle était composé d'une majorité de filles. Toutefois, ceci représente bien les ratios de genre habituellement observés chez les enfants ayant des troubles spécifiques d'apprentissage en lecture, la dyslexie étant plus fréquente chez les garçons (ratio allant d'environ 2/1 à 3/4; American Psychiatric Association, 2015; Rutter et al., 2004). De plus, comme les scores des filles aux six mesures du TELEQ ne différaient pas significativement de ceux des garçons, la non-équivalence des groupes sur cette variable n'apparaît pas problématique. Ensuite, il y avait davantage de mères ayant un diplôme universitaire (1^{er}, 2^e ou 3^e cycle) dans le groupe contrôle que dans le groupe clinique. Par ailleurs, les enfants du groupe contrôle dont la mère détenait un diplôme universitaire obtenaient en moyenne des scores plus élevés à quatre des six mesures du TELEQ que ceux dont la mère ne détenait pas de diplôme universitaire. Ces effets étaient toutefois de petite ou moyenne taille. Ceci reflète une observation souvent rapportée dans la littérature, soit que le niveau de scolarité des parents influence le développement des habiletés de lecture de l'enfant (p. ex. Écalle et Magnan, 2015; Ministère de l'Éducation, 2003). Toutefois, puisque le niveau d'éducation de la mère semble associé à certains scores du TELEQ chez l'enfant, il est possible que cette variable ait influencé les différences de groupes. Finalement, les mesures d'estimation du quotient intellectuel verbal et non verbal étaient significativement plus élevées dans le groupe contrôle que dans le groupe clinique. De plus, certaines

mesures de langage écrit du TELEQ étaient faiblement ou moyennement corrélées à l'estimation de quotient intellectuel des enfants, tel que souvent rapporté dans la littérature (p. ex. Torgesen, 1989). Notons toutefois qu'autant pour l'estimation du quotient intellectuel verbal que non verbal, les moyennes des deux groupes se trouvaient dans la moyenne. Les mesures d'estimation du quotient intellectuel n'ont pas été contrôlées dans les analyses puisque les groupes d'enfants ayant des troubles neurodéveloppementaux, comme la dyslexie, obtiennent en général des moyennes plus faibles aux mesures de quotient intellectuel que les groupes d'enfants contrôle (p. ex. via l'atteinte de la mémoire de travail verbale ou du vocabulaire qui sont également impliqués dans les mesures de quotient intellectuel ou encore, car le quotient intellectuel est affecté négativement par les mêmes facteurs étiologiques que les déficits en lecture et écriture; de Jong et van Bergen, 2017). Ainsi, il n'est pas recommandé d'utiliser le quotient intellectuel comme covariable, car ceci aurait pour conséquence de retirer toute la variabilité liée à cette association (Dennis et al., 2009).

En deuxième lieu, le faible échantillon de participants de 2^e année ayant reçu un diagnostic de dyslexie/dysorthographe représente une limite puisque ce niveau scolaire n'a pu être inclus dans les comparaisons de groupe. Puisqu'il est difficile d'établir formellement la présence d'une dyslexie avant l'âge de 7 ans (Harrison, 2007), il était plus ardu de recruter des participants ayant déjà reçu leur diagnostic en 2^e année. Cela peut s'expliquer par le fait que les cliniciens doivent s'assurer que les difficultés observées chez un enfant persistent malgré une intervention ciblée, appropriée et d'intensité suffisante pour pouvoir distinguer un retard (causé par un enseignement formel inadéquat et qui pourra se rattraper) d'un trouble (provenant de difficultés intrinsèques à l'enfant; St-Pierre et al., 2010). Cet élément de résistance à l'intervention est d'ailleurs mis de l'avant comme condition essentielle au diagnostic de dyslexie-dysorthographe par plusieurs auteurs (Fletcher et al., 2005; Fuchs et al., 2002).

En troisième lieu, une autre limite concerne le fait que les enfants composant le groupe clinique avaient reçu leur diagnostic en moyenne entre 1 et 2 ans avant l'évaluation. Durant ces années, ils ont pu bénéficier d'interventions orthopédagogiques ou orthophoniques visant l'amélioration de leurs habiletés de décodage ou d'orthographe. Ceci représente une limite en ce sens où il est possible que les enfants du groupe clinique aient obtenu de meilleurs scores au TELEQ en raison des interventions reçues que s'ils avaient effectué le TELEQ au moment de leur diagnostic, diminuant ainsi la sensibilité apparente du TELEQ ou

baisant les valeurs seuil optimales à utiliser auprès d'enfants qui n'ont pas encore reçu leur diagnostic. De futures études évaluant la sensibilité et la spécificité de l'outil pourraient être réalisées auprès d'enfants venant tout juste de recevoir leur diagnostic de dyslexie/dysorthographe, ce qui permettra également d'inclure plus d'enfants de la 2^e année primaire présentant une dyslexie/dysorthographe.

En quatrième lieu, des participants ont été exclus du groupe contrôle sur la base d'un trouble neurodéveloppemental connu ou fortement suspecté (scores déficitaires aux estimations du QI, à l'*Alouette-R* ou aux symptômes d'inattention ou d'hyperactivité), afin que ce groupe soit constitué d'enfants ayant un développement typique. Ces exclusions dans l'échantillon normatif peuvent rendre les seuils du TELEQ plus sévères et faire en sorte d'identifier des enfants faibles lecteurs comme ayant des difficultés de lecture ou empêcher la catégorisation du degré de sévérité des difficultés (McFadden, 1996). Lors de l'augmentation de la taille de l'échantillon normatif, il pourra être intéressant de réintégrer un certain pourcentage d'enfants ayant un trouble neurodéveloppemental pour représenter l'étendue complète des scores au TELEQ dans la population générale (p. ex. méthode utilisée pour les normes du WISC-V; Wechsler, 2015).

Finalement, l'absence de mots réguliers dans les deux sous-tests du TELEQ représente un choix intéressant pour réduire le temps de passation et ainsi augmenter l'efficacité des évaluations. La comparaison des scores pondérés obtenus aux deux listes de mots du TELEQ est suffisante pour permettre l'évaluation de l'intégrité des voies phonologique et lexicale. Toutefois, inclure dans le TELEQ une liste de lecture de mots réguliers fréquents, peu fréquents et de pseudomots appariés aurait permis au clinicien, en comparant leur vitesse de lecture, d'avoir un élément supplémentaire pour se prononcer sur la voie privilégiée par l'enfant (St-Pierre et al., 2010).

Conclusion

En conclusion, les sous-tests « Lecture de mots et de pseudomots » et « Dictée de mots et de pseudomots » du TELEQ permettent de discriminer adéquatement les enfants ayant une dyslexie/dysorthographe de ceux n'en ayant pas. Ceci indique qu'ils peuvent être utilisés dans le cadre d'une démarche diagnostique chez les enfants de la 2^e à la 6^e année primaire fréquentant les écoles francophones du Québec. Les valeurs seuils optimales et les normes préliminaires contenues dans cet article pourront être utilisées par les cliniciens habilités à évaluer les troubles de langage écrit afin de les aiguiller dans leurs diagnostics. D'autres sous-tests du TELEQ sont en cours

de développement ou de validation. À terme, le TELEQ comprendra différents sous-tests adaptés à la clientèle franco-québécoise permettant l'évaluation des habiletés de langage écrit.

Références

- Aguert, M. et Capel, A. (2018). Mieux comprendre les scores z pour bien les utiliser. *Rééducation orthophonique*, 55(274), 61-86. https://bib.vinci.be/opac_css/index.php?lvl=notice_display&id=214151
- American Psychiatric Association. (2015). *DSM-5: manuel diagnostique et statistique des troubles mentaux* (5^e éd.; traduit par M.-A. Crocq et J.-D. Guelfi). Elsevier Masson.
- American Speech-Language-Hearing Association. (2001). *Roles and responsibilities of speech-language pathologists with respect to reading and writing in children and adolescents* [Position Statement]. <https://www.asha.org/policy/ps2001-00104/>
- Baudry, M.-È., Laniel, P., Malo-Véronneau, L., Picotte-Lavoie, M. et Gauthier, B. (2020). TELEQ : création et pré-validation d'un outil québécois d'évaluation de l'orthographe. *Revue canadienne d'orthophonie et d'audiologie*, 44(2), 87-106. https://cjslpa.ca/files/2020_CJSLPA_Vol_44/No_2/CJSLPA_Vol_44_No_2_2020_MS_1181.pdf
- Baudry, M.-È., Laniel, P. et Gauthier, B. (2021). *Test d'évaluation du langage écrit québécois: Manuel technique et guide d'administration—Dictée de mots et de pseudomots*. (version 1.0). Université de Montréal. <https://teleq.ca/telechargerleteleq.html>
- Benjamini, Y. et Hochberg, Y. (2000). On the adaptive control of the false discovery rate in multiple testing with independent statistics. *Journal of Educational and Behavioral Statistics*, 25(1), 60-83. <https://doi.org/10.3102/10769986025001060>
- Bertrand, D., Fluss, J., Billard, C. et Ziegler, J. C. (2010). Efficacité, sensibilité, spécificité : comparaison de différents tests de lecture. *L'Année psychologique*, 110(2), 299-320. <https://doi.org/10.4074/S000350331000206X>
- Bouchard, M.-È. G., Fitzpatrick, E. M. et Olds, J. (2009). Analyse psychométrique d'outils d'évaluation utilisés auprès des enfants francophones. *Revue canadienne d'orthophonie et d'audiologie*, 33(3), 129-139. https://cjslpa.ca/files/2009_CJSLPA_Vol_33/No_03_113-160/Bouchard_Fitzpatrick_Olds_CJSLPA_2009.pdf
- Bujang, M. A., Sa'at, N., Sidik, Tg Abu Bakar Sidik, T. M. I. et Lim, C. J. (2018). Sample size guidelines for logistic regression from observational studies with large population: Emphasis on the accuracy between statistics and parameters based on real life clinical data. *The Malaysian Journal of Medical Sciences*, 25(4), 122-130. <https://doi.org/10.21315/mjms2018.25.4.12>
- Casalis, S. (2005). Dyslexies du développement et reconnaissance visuelle de mots : questions introductives. Dans Y. Coello, C. Moroni et S. Casalis (dir.), *Vision, espace et cognition : Fonctionnement normal et pathologique* (p. 165-179). Presses universitaires du Septentrion.
- Castles, A. et Coltheart, M. (1993). Varieties of developmental dyslexia. *Cognition*, 47(2), 149-180. [https://doi.org/10.1016/0010-0277\(93\)90003-E](https://doi.org/10.1016/0010-0277(93)90003-E)
- Cavalli, E. et Colé, P. (2018). Screening for dyslexia in french-speaking university students: An evaluation of the detection accuracy of the Alouette test. *Journal of Learning Disabilities*, 51(3), 268-282. <https://doi.org/10.1177/0022219417704637>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Routledge.
- Coltheart, M., Rastle, K., Perry, C., Langdon, R. et Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, 108(1), 204-256. <http://dx.doi.org/10.1037/0033-295X.108.1.204>
- Conners, K. C. (2008). *Conners* (3^e éd.) Multi-Health Systems.
- Curran, P. J., West, S. G. et Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, 1(1), 16-29. <https://doi.org/10.1037/1082-989X.1.1.16>
- Daniel, S. S., Walsh, A. K., Goldston, D. B., Arnold, E. M., Reboussin, B. A. et Wood, F. B. (2006). Suicidality, school dropout, and reading problems among adolescents. *Journal of Learning Disabilities*, 39(6), 507-514. <https://doi.org/10.1177/00222194060390060301>

- de Jong, P. F. et van Bergen, E. (2017). Issues in diagnosing dyslexia. Dans E. Segers et P. van den Broek (dir.), *Developmental perspectives in written language and literacy* (p. 349-361). John Benjamins. <https://doi.org/10.1075/z.206.21dej>
- Delacour, H., Servonnet, A. et Roche, C. (2009). Critères d'évaluation de la validité d'un test biologique. *Revue Francophone des Laboratoires*, 2009(412), 41-48. [https://doi.org/10.1016/S1773-035X\(09\)73934-6](https://doi.org/10.1016/S1773-035X(09)73934-6)
- Dennis, M., Francis, D. J., Cirino, P. T., Schachar, R., Barnes, M. A. et Fletcher, J. M. (2009). Why IQ is not a covariate in cognitive studies of neurodevelopmental disorders. *Journal of the International Neuropsychological Society*, 15(3), 331-343. <https://doi.org/10.1017/S1355617709090481>
- DuPaul, G. J., Power, T. J., Anastopoulos, A. D. et Reid, R. (1998). *ADHD Rating Scale-IV: Checklists, norms, and clinical interpretation*. Guilford Press.
- Écalle, J. et Magnan, A. (2015). *L'apprentissage de la lecture et ses difficultés*. Dunod. <https://doi.org/10.3917/dunod.ecall.2015.01>
- Fletcher, J. M., Francis, D. J., Morris, R. D. et Lyon, G. R. (2005). Evidence-based assessment of learning disabilities in children and adolescents. *Journal of Clinical Child & Adolescent Psychology*, 34(3), 506-522. https://doi.org/10.1207/s15374424jccp3403_7
- Friberg, J. C. (2010). Considerations for test selection: How do validity and reliability impact diagnostic decisions? *Child Language Teaching and Therapy*, 26(1), 77-92. <https://doi.org/10.1177/0265659009349972>
- Fuchs, L. S., Fuchs, D. et Speece, D. L. (2002). Treatment validity as a unifying construct for identifying learning disabilities. *Learning Disability Quarterly*, 25(1), 33-45. <https://doi.org/10.2307/1511189>
- Glaros, A. G. et Kline, R. B. (1988). Understanding the accuracy of tests with cutting scores: The sensitivity, specificity, and predictive value model. *Journal of Clinical Psychology*, 44(6), 1013-1023. [https://doi.org/10.1002/1097-4679\(198811\)44:6<1013::AID-JCLP2270440627>3.0.CO;2-Z](https://doi.org/10.1002/1097-4679(198811)44:6<1013::AID-JCLP2270440627>3.0.CO;2-Z)
- Glascio, F. P. et Byrne, K. E. (1993). The accuracy of three developmental screening tests. *Journal of Early Intervention*, 17(4), 368-379. <https://doi.org/10.1177/105381519301700403>
- Hajjan-Tilaki, K. (2013). Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. *Caspian Journal of Internal Medicine*, 4(2), 627-635.
- Harrison, A. G. (2007). Recommended best practices for the early identification and diagnosis of children with specific learning disabilities in Ontario. *Canadian Journal of School Psychology*, 20(1-2), 21-43. <https://doi.org/10.1177/0829573506295461>
- Hosmer, D. W. et Lemeshow, S. (2000). Assessing the fit of the model. Dans W. A. Shewhart, et S. S. Wilks (dir.), *Applied logistic regression* (2^e éd., p. 143-202). John Wiley and Sons Inc. <https://doi.org/10.1002/0471722146.ch5>
- Humphrey, N. et Mullins, P. M. (2002). Self-concept and self-esteem in developmental dyslexia. *Journal of Research in Special Educational Needs*, 2(2). <https://doi.org/10.1111/j.1471-3802.2002.00163.x>
- Institut national de la santé et de la recherche médicale. (2007). *Dyslexia dysorthography dyscalculia: Review of the scientific data*. <https://www.ncbi.nlm.nih.gov/books/NBK10787/>
- International Dyslexia Association. (2002). *Definition of dyslexia*. <https://dyslexiaida.org/definition-of-dyslexia/>
- Jacquier-Roux, M., Lequette, C., Pouget, G., Valdois, S. et Zorman, M. (2010). *Batterie analytique du langage écrit*. Groupe Cogni-Sciences. <http://www.cognisciences.com/accueil/outils/article/bale>
- Lalkhen, A. G. et McCluskey, A. (2008). Clinical tests: Sensitivity and specificity. *Continuing Education in Anaesthesia Critical Care & Pain*, 8(6), 221-223. <https://doi.org/10.1093/bjaceaccp/mkn041>
- Laniel, P., Vallières-Lavoie, G., Champagne, L. et Gauthier, B. (2022). Création et prévalidation du sous-test Lecture de mots et de pseudomots du TELEQ. *Revue canadienne d'orthophonie et d'audiologie*, 46(2), 141-159.
- Leclercq, A.-L. et Veys, E. (2014). Réflexions sur le choix de tests standardisés lors du diagnostic de dysphasie. *Approche neuropsychologique des apprentissages chez l'enfant*, 26(131-4), 374-382.
- Lefavrais, P. (2005). *Alouette-R : Test d'analyse de la lecture et de la dyslexie*. Édition du Centre de Psychologie Appliquée.
- Lefebvre, P. et Stanké, B. (2016). Les dyslexies-dysorthographies développementales. Dans B. Stanké (dir.), *Les dyslexies-dysorthographies* (1^{re} éd., p. 47-68). Presses de l'Université du Québec. <https://doi.org/10.2307/j.ctv1rr6d74.8>
- Lefebvre, P. et Trudeau, N. (2005). L'orthophoniste et les tests normalisés. *Fréquences*, 17(2), 17-20.
- Livingston, E. M., Siegel, L. S. et Ribary, U. (2018). Developmental dyslexia: Emotional impact and consequences. *Australian Journal of Learning Difficulties*, 23(2), 107-135. <https://doi.org/10.1080/19404158.2018.1479975>
- Lyon, G. R., Shaywitz, S. E. et Shaywitz, B. A. (2003). A definition of dyslexia. *Annals of Dyslexia*, 53(1), 1-14. <https://doi.org/10.1007/s11881-003-0001-9>
- Margari, L., Buttiglione, M., Craig, F., Cristella, A., de Giambattista, C., Matera, E., Operto, F. et Simone, M. (2013). Neuropsychopathological comorbidities in learning disorders. *BMC Neurology*, 13(198), 1-698. <https://doi.org/10.1186/1471-2377-13-198>
- McCauley, R. J. et Swisher, L. (1984). Psychometric review of language and articulation tests for preschool children. *Journal of Speech and Hearing Disorders*, 49(1), 34-42. <https://doi.org/10.1044/jshd.4901.34>
- McFadden, T. U. (1996). Creating language impairments in typically achieving children. *Language, Speech, and Hearing Services in Schools*, 27(1), 3-9. <https://doi.org/10.1044/0161-1461.2701.03>
- Ministère de l'Éducation. (2003). *Les difficultés d'apprentissage à l'école: cadre de référence pour guider l'intervention*. <http://www.education.gouv.qc.ca/references/tx-solrtypercherchepublicationtx-solrpublicationnouveauite/resultats-de-la-recherche/detail/article/les-difficultes-dapprentissage-a-lecole/>
- Ministère de l'Éducation, du Loisir et du Sport. (2014). *Programme de formation de l'école québécoise, enseignement primaire: Liste orthographique à l'usage des enseignantes et des enseignants*. http://www.education.gouv.qc.ca/fileadmin/site_web/documents/education/jeunes/pfeq/Liste-orthographique-document-reference.pdf
- Ministère de l'Éducation et de l'Enseignement supérieur. (2020). *Indices de défavorisation 2017-2018*. <http://www.education.gouv.qc.ca/references/tx-solrtypercherchepublicationtx-solrpublicationnouveauite/resultats-de-la-recherche/detail/article/indices-de-defavorisation/>
- Morin, V., Morin, J. F., Mercier, M., Moineau, M. P. et Codet, J. P. (1998). Les courbes ROC en biologie médicale; Clock around the ROC. *Immuno-analyse & biologie spécialisée*, 13(5), 279-286. [https://doi.org/10.1016/S0923-2532\(98\)80016-1](https://doi.org/10.1016/S0923-2532(98)80016-1)
- Mousty, P. et Alegria, J. (1999). L'acquisition de l'orthographe: données comparatives entre enfants normo-lecteurs et dyslexiques. *Revue française de pédagogie*, 126(1), 7-22. <https://doi.org/10.3406/rfp.1999.1091>
- Perkins, N. J. et Schisterman, E. F. (2006). The inconsistency of "optimal" cutpoints obtained using two criteria based on the Receiver Operating Characteristic Curve. *American Journal of Epidemiology*, 163(7), 670-675. <https://doi.org/10.1093/aje/kwj063>
- Picotte-Lavoie, M. (2021). *TELEQ: création et pré-validation d'un test de compréhension de lecture pour les enfants québécois de niveau scolaire primaire* [essai doctoral, Université de Montréal]. Papyrus. <http://hdl.handle.net/1866/25703>
- Plante, E. et Vance, R. (1994). Selection of preschool language tests: A data-based approach. *Language, Speech, and Hearing Services in Schools*, 25(1), 15-24. <https://doi.org/10.1044/0161-1461.2501.15>
- Pourcin, L., Sprenger-Charolles, L., El Ahmadi, A. et Colé, P. (2016). Reading and related skills in Grades 6, 7, 8 and 9: French normative data from EVALEC. *European Review of Applied Psychology*, 66(1), 23-37. <https://doi.org/10.1016/j.erap.2015.11.002>
- Prior, M., Smart, D., Sanson, A. et Oberklaid, F. (1999). Relationships between learning difficulties and psychological problems in preadolescent children from a longitudinal sample. *Journal of the American Academy of Child & Adolescent Psychiatry*, 38(4), 429-436. <https://doi.org/10.1097/00004583-199904000-00016>
- Rousseau, N. (2016). L'apprentissage et la persévérance scolaires des élèves ayant des troubles d'apprentissage: Une perspective psychopédagogique de l'intervention. Dans B. Stanké (dir.), *Les dyslexies-dysorthographies* (1^{re} éd., p. 197-214). Presses de l'Université du Québec. <https://doi.org/10.2307/j.ctv1rr6d74.13>

- Rutter, M., Caspi, A., Fergusson, D., Horwood, L. J., Goodman, R., Maughan, B., Moffitt, T. E., Meltzer, H. et Carroll, J. (2004). Sex differences in developmental reading disability: New findings from 4 epidemiological studies. *Journal of the American Medical Association*, 291(16), 2007-2012. <https://doi.org/10.1001/jama.291.16.2007>
- Shaywitz, S. E., Fletcher, J. M. et Shaywitz, B. A. (1994). Issues in the definition and classification of attention deficit disorder. *Topics in Language Disorders*, 14(4), 1-25. <https://doi.org/10.1097/00011363-199408000-00003>
- Siegel, L. S. et Hurford, D. P. (2019). The case against discrepancy models in the evaluation of dyslexia. *Perspectives on Language and Literacy*, 45(1), 23-28. <https://dyslexialibrary.org/wp-content/uploads/file-manager/public/1/Winter%202019%20Final%20Siegel%20Hurford%20p23-28.pdf>
- Spaulding, T. J., Plante, E. et Farinella, K. A. (2006). Eligibility criteria for language impairment. *Language, Speech, and Hearing Services in Schools*, 37(1), 61-72. [https://doi.org/10.1044/0161-1461\(2006\)007](https://doi.org/10.1044/0161-1461(2006)007)
- Sprenger-Charolles, L., Lacert, P., Béchenec, D., Colé, P. et Serniclaes, W. (2001). Stabilité dans le temps et inter-langues des sous-types de dyslexie développementale. *Approches neuropsychologiques des apprentissages chez l'enfant*, 62-63, 115-128.
- Stanké, B. et Lefebvre, P. (2016). La dyslexie-dysorthographe phonologique. Dans B. Stanké (dir.), *Les dyslexies-dysorthographies* (1^{re} éd., p. 69-102). Presses de l'Université du Québec. <https://doi.org/10.2307/j.ctv1rr6d74.9>
- St-Pierre, M.-C., Dalpé, V., Lefebvre, P. et Giroux, C. (2010). *Difficultés de lecture et d'écriture : Prévention et évaluation orthophonique auprès des jeunes*. (1^{re} éd.) Presses de l'Université du Québec. <https://doi.org/10.2307/j.ctv18pgjzc>
- Torgesen, J. K. (1989). Why IQ is relevant to the definition of learning disabilities. *Journal of Learning Disabilities*, 22(8), 484-486. <https://doi.org/10.1177/002221948902200806>
- Vallières-Lavoie, G., Laniel, P. et Gauthier, B. (2021) *Test d'évaluation du langage écrit québécois: Manuel technique et guide d'administration – Lecture de mots et de pseudomots*. (version 1.0). Université de Montréal. <https://teleq.ca/telechargerleteleq.html>
- Wechsler, D. (2005). *Échelle d'intelligence de Wechsler pour enfants – Quatrième édition – Version pour francophones du Canada*. Pearson Canada Assessment.
- Wechsler, D. (2015). *Échelle d'intelligence de Wechsler pour enfants – Cinquième édition : Version pour francophones du Canada*. Pearson Canada Assessment.
- Willcutt, E. G. et Pennington, B. F. (2000). Comorbidity of reading disability and attention-deficit/hyperactivity disorder: Differences by gender and subtype. *Journal of Learning Disabilities*, 33(2), 179-191. <https://doi.org/10.1177/002221940003300206>
- Youden, W. J. (1950). Index for rating diagnostic tests. *Cancer*, 3(1), 32-35. [https://doi.org/10.1002/1097-0142\(1950\)3:1<32::AID-CNCR2820030106>3.0.CO;2-3](https://doi.org/10.1002/1097-0142(1950)3:1<32::AID-CNCR2820030106>3.0.CO;2-3)

Note des auteurs

Les demandes au sujet de cet article doivent être acheminées à Patricia Laniel, Université de Montréal, 1700, rue Jacques-Tétreault, Laval, QC, Canada, H7N 0A5. Courriel : patricia.laniel@umontreal.ca

Remerciements

Cette recherche a été financée par une subvention institutionnelle du CRSH-Université de Montréal (Programme subvention d'exploration) et par le Fonds d'installation de nouveaux professeurs de l'Université de Montréal au dernier auteur. Merci aux Fonds de recherche du Québec – Société et culture pour la bourse de doctorat en recherche octroyée à la première auteure.

Remerciements particuliers à Diane Jacques et Marina Attié pour leur contribution au développement du TELEQ.

Déclaration

Les auteurs déclarent n'avoir aucun conflit d'intérêts, financiers ou autres.